Comprehension, Use Cases and Requirements

Keith Phalp, Anita Adlem, Sheridan Jeary, Jonathan Vincent

Software Systems Research Centre,
Bournemouth University, Fern Barrow, Poole,
Dorset, BH12 5BB
kphalp, aadlem, sjeary, jvincent@bournemouth.ac.uk

Abstract

Within requirements engineering it is generally accepted that in writing specifications (or indeed any requirements phase document), one attempts to produce an artefact which will be simple to comprehend for the user. That is, whether the document is intended for customers to validate requirements, or engineers to understand what the design must deliver, comprehension is an important goal for the author. Indeed, advice on producing 'readable' or 'understandable' documents is often included in courses on requirements engineering. However, few researchers, particularly within the software engineering domain, have attempted either to define or to understand the nature of comprehension and it's implications for guidance on the production of quality requirements.

In contrast, this paper examines thoroughly the nature of textual comprehension, drawing heavily from research in discourse process, and suggests some implications for requirements (and other) software documentation. In essence, we find that the guidance on writing requirements, often prevalent within software engineering, may be based upon assumptions which are an oversimplification of the nature of comprehension. Furthermore, that these assumptions may lead to rules which detract from the quality of the requirements document and, thus, the understanding gained by the reader. Finally the paper suggests lessons learned which may be useful in formulating future guidance for the production of requirements documentation.

1.0 Introduction: The importance of Comprehension

The complete, accurate and concise documenting of requirements is of vital, perhaps paramount importance [1] within software development, and errors made in this phase are often considered the most difficult to solve and most costly to fix [2]. Hence, the potential benefits of successful comprehension [3], promise improvements in software quality, stakeholder satisfaction and development costs. However, comprehension is 'a complex interaction of basic cognitive processes' [4] which can be regarded as 'one of the most complex and uniquely human of

cognitive activities' [5], and thus, understanding the nature of comprehension is far from trivial. .

One area which offers a valid means of investigating comprehension is that of Discourse Process. Bamberg and Moissinac [6] define discourse as 'broadly taken to mean the use of language beyond that of a single sentence'. Discourse Process analyses the way in which sequences of sentences combine to produce coherent sections of language; and thus extends the traditional linguistic study of the construction of individual sentences [7, 8]. Requirements techniques aim to communicate sets of concepts and meanings built, often within sizeable documents, to various stakeholders in a system. Moreover, they typically capture the interaction between an actor and the system in order to accomplish the actor's goal across a number of statements [9, 10]. Therefore, discourse process is particularly relevant for the understanding, and improvement, of requirements.

Hence, the aim of this paper is to investigate the lessons that can be learned from the field of Discourse Process, particularly about the nature of comprehension, for improving the writing of requirements and specifications. Further work, outside the scope of this paper, examines the implication of these ideas for existing requirements authoring guidelines. However, our intention here is, by thorough scrutiny of discourse process, to highlight the issues and provide lessons learned. The following section (two) introduces principles of discourse process, section three describes application of discourse process theory, sections four and five discuss the roles of background knowledge and of the reader, section six describes limitations of the approach, and, finally, section seven offers some conclusions.

2.0 General Principles of Discourse Process

As people read and comprehend, they build a mental representation or model of the text and of the situations portrayed within the text. This mental representation consists of multiple levels of code in memory. Of the various levels of representation that have been proposed, most discourse psychologists accept three: surface code, textbase, and situation model. The surface code consists of the exact wording and grammar of clauses, i.e. it preserves the actual text. The textbase preserves the meaning of the explicitly stated information in the text but not its exact wording and grammar. It is usually viewed as consisting of a structured set of text propositions. Each proposition refers to a state, event, or action and may possess a truth value. It contains a predicate (e.g. a main verb, adverb, adjective or connective) and one or more arguments (e.g. nouns or other propositions embedded in it). The predicate denotes a relation between arguments. Each argument plays a functional role, such as agent (the entity which causes the action of the verb), patient (the entity which undergoes the action of the verb), instrument or location.

The propositions form a network, being linked to each other through shared arguments and other connectives in the text. They typically form hierarchical structures. As more text is read, new propositions are added to the network,

forming a coherent knowledge structure. The textbase represents a superficial understanding and enables such tasks as recall of the text or the answering of questions which only require information from one sentence.

The situation model is constructed from inferences (additional propositions [11]) made by the reader as they extract the meaning from the text and activate relevant background world knowledge from their memory stores. Almost every aspect of comprehension depends at least partly on these inferences constructed by the reader. These are necessary for filling in gaps and thus building the situation model and for increasing coherence. There are two types, bridging inferences and elaborative inferences. Bridging inferences link the current clause to the preceding text, thus maintaining local text coherence. Elaborative inferences are ideas which are strongly implied by the context of a discourse; they supply additional propositions based on world knowledge and elaborate the text which is being read, but do not affect coherence [12]. The situation model, therefore, represents deeper understanding and enables the connecting of two or more sentences of the text, the linking of information from the text with background knowledge, and problem solving. All of these levels contribute to the meaning representations that readers mentally build as they comprehend.

2.1 Coherence

If a text is to be comprehensible, it must be coherent, i.e. the reader must be able to identify relations among the ideas in the text [12]. For example, many researchers have argued that the presence of expressions in a text referring to the same entities (coreference) promotes coherence [13]. Two main types of coherence exist: local and global. Local coherence occurs when each sentence is related to the next, e.g. through argument overlap (i.e., the incoming statement of a text contains a noun or pronoun which refers to the same entity as a noun or pronoun in the preceding sentence). Global coherence requires that larger sections of text are clearly related to each other and to the overall topic. For Graesser et al. [8] global coherence is achieved if the reader can connect the incoming statement with the overall structure of the text or with information from the text which is no longer held in working memory. Gibbs [14], however, states that cognitive psychologists have argued that the comprehension of text is driven by the reader's search for cause and effect and that these causal relations provide much of the global coherence of memory. Graesser et al. [8] similarly take the view that readers actively seek causal explanations of why events in the situation model occur and why the writer conveys information as they are reading. McNamara et al. [15] also refer to explanatory coherence which is achieved when the content of the text supplies background knowledge which the reader may lack but which is needed to understand the text. Studies have demonstrated that increases in the local, global and explanatory coherence of texts have resulted in substantial increases in text recall by readers [16].

2.2 Advantage of First Mention

The Advantage of First Mention proposes that the first entity mentioned in a clause, regardless of grammatical position or semantic role, enjoys a higher level of

activation and is accessed more quickly from memory than the other entities. This is explained by the Structure Building Framework [17] which is described below.

2.3 Structure Building Framework

According to the Structure Building Framework, the aim of comprehension is to create cohesive mental representations, or structures. When readers build a mental representation of the text, they perform three processes. First, they lay a foundation when a new topic is introduced by setting up the mental structures needed to establish coherence. Aaronson and Ferres (cited by Foltz [11] and Gernsbacher [17]) argue that readers spend more time processing the first word of a clause or sentence. This is attributed to the time taken to generate an initial framework for processing the new discourse. They then build upon this foundation, if possible, by mapping any information from the subsequent text which is relevant to the same topic onto the structures that have been developed. If the subsequent text is not relevant, however, the reader shifts attention and builds a new substructure. These processes produce a branching structure which is an overall cohesive mental representation of the text.

2.4 Event Indexing Model

In Zwaan et al.'s Event Indexing Model [18] the reader of simple stories, or literary short stories, monitors five conceptual dimensions in order to construct a multithreaded situation model. The five dimensions are: protagonists, causality, protagonists' goals, geographic space and time. When the reader processes a new statement, a break in continuity may occur on any of these dimensions. Zwaan et al. [18] found that the time taken to read a particular event in a story increased as a function of the number of these dimensions which contained a discontinuity (i.e., lacked coherence). Each dimension had a unique impact on reading time. Therefore, the Event Indexing Model recommends that the writer should strive, wherever possible, to avoid a break in continuity on any of these dimensions.

2.5 NVN Strategy

The NVN (Noun Verb Noun) strategy is one of the most important heuristics in comprehension [19]. The reader is inclined to assume that the subject of a sentence is also the agent of some action, and this can be seen in much guidance (such as for use case descriptions) which suggests adopting, subject, verb, object structures.

3.0 Application of Discourse Process Theories

3.1 Pronouns

There are various discourse process theories regarding pronouns. These theories include: Gernsbacher's Advantage of First Mention, "centering theory" (discussed below), the subject assignment strategy, which states that if a pronoun is ambiguous the reader is likely to interpret it as referring to the subject of the preceding sentence [20], and the theories of a) Matthews and Chodorow and b) Sidner (both cited by Greene et al. [21]) who argue that an entity which is the grammatical subject of a sentence is a strong candidate for identification with a subsequent pronoun which is also in the subject position. If these theories are

accepted, then the success of a pronoun depends not just upon the pronoun itself but also on the context of its referent. Therefore, the writer who wishes to reduce the risk of ambiguity associated with a pronoun must take care to ensure that the referent of the pronoun is introduced in the subject position.

Certain discourse process theories support the use of pronouns, however. These include the theory that the processes by which the referent of a pronoun is identified actually increase the accessibility in memory of the intended referent [21]. Once again, opposing theories exist; Gernsbacher [17], for example, argues that all anaphoric references increase accessibility. McKoon et al. [22] also reject the theory that pronouns increase accessibility, proposing instead the theory of memory-based text processing, which is based on the principle that general memory processes make the intended referents available to the reader even before the pronouns themselves are used.

The use of pronouns can sometimes avoid awkwardness, but this is a factor which is largely based on intuition rather than empirical evidence. Tanenhaus et al. and Tanenhaus and Carlson (both cited in Gordon and Chan [23]) demonstrate that certain kinds of references to entities seem to be understood as a result of the entities and references sharing the same thematic roles across sentences. In such circumstances, an explicit reference to the entity, e.g., repeating the noun rather than using a pronoun, results in a text which is awkward to read. Thus, there is a trade-off between avoiding awkwardness and its accompanying increase in reading time on the one hand and avoiding ambiguity on the other. However, increases in reading time are of minor importance, often being measured in milliseconds, while the importance of avoiding the risk of ambiguity is paramount, and particularly important within our context of requirements and specification.

In summary the writer should consider carefully where to use pronouns. There is a trade-off between the benefits of increased coherence on the one hand and the increased risk of ambiguity coupled with increased time and effort on the other. Furthermore, although Gordon et al. [24] demonstrate that pronouns can affect reading time, they also indicate that the use of pronouns has no significant effects on readers' accuracy when answering comprehension questions. This has clear implications for requirements engineering, where accuracy of interpretation far outweighs minor differences in the speed of comprehension. Pragmatically, it would therefore be wiser for the writer to avoid using them, given the ambiguity that they can create, with little apparent gain for the reader.

3.2 Use of the Passive Voice

The disadvantages of using the passive voice are demonstrated by many researchers in the fields of linguistics and discourse process. A number of researchers demonstrate (cited by Gordon and Chan[23]) that when sentences are presented in isolation, passive versions that can be changed into active voice without losing their meaning take longer to understand than their active form. Similarly, the results of an experiment performed by Gordon and Chan [23]

showed that passive sentences were read more slowly than active sentences which had a similar meaning. Indeed, the increase in time taken is not just limited to reading and understanding sentences; experiments performed by Ferreira [19] demonstrate that passive sentences also take longer to formulate than active ones. Hence, there is a cost for the author as well as the reader where the passive is concerned.

Of greater importance than the speed of writing and reading is the accuracy of understanding. Ferreira [19] states that evidence suggests that syntactic complexity increases the chances of misinterpreting text. Her own experiments indicate that comprehenders performing the task of identifying agents and patients in active and passive versions of the same sentences not only take longer but are less accurate when answering questions about the passive forms. Indeed, Ferreira's findings consistently demonstrate that the passive form is more difficult to understand than its active counterpart. The results of her experiments indicate that even nonreversible passive sentences may not be particularly easy to understand, although previous researchers have claimed that this is the case [25]. (A nonreversible sentence is one in which the subject and object cannot be transposed without the sentence becoming implausible.)

The major benefit of the passive is that it allows an entity which is not the thematic agent to be the grammatical subject of a sentence. The importance of the grammatical subject and of the first mentioned entity in a sentence is stressed by various theories and studies in discourse process and linguistics. Clark (cited by Johnson-Laird [26]) claims, for example, that "people put what they want to talk about…in the beginning of the sentence" and Ferreira [27] notes that there was a trend for participants in her tests to be more accurate with the first role that they encountered in a sentence.

In summary, for the context of requirements documents, it appears that the increased risk of misinterpretation found in use of the passive voice, would suggest that avoidance of the passive appears to be the safest option.

3.3 Simplicity

Where the aim is to increase the ease of comprehension, it seems sensible to avoid the use of complex structures, such as the passive voice, the negative form, and adjectives and adverbs. Moreover, this is supported by discourse process theories and studies which indicate that complexity increases reading time (e.g., [28] and [19]) and that challenging structures are sometimes misinterpreted. However, even the concept of simplicity is complicated; for example, Ferreira [27] notes that studies exist which demonstrate that comprehenders are generally willing to accept wildly implausible meanings of simple sentences. Moreover, good use of discourse cues, e.g., words or phrases, such as "because", "first", "although" and "also", which signal grammatical and semantic relationships between entities in a discourse, improves comprehension and recall [12; although, as a caveat, indiscriminate use of semantic cues may have damaging effects on the reader's

ability to recall the text [12]. Evidence suggests that how beneficial discourse cues are to readers may depend upon their reading skills or level of prior domain knowledge [29]. The situation is further complicated by the problem that whether a concept is new or not is not just dependent on the preceding text; it is also dependent on the reader's knowledge, that is, what is new to one reader may not be new to another.

Thus, there is a trade-off between simple sentences which can result in reduced reading time and decreased risk of misinterpretation and more complicated structures which can also lead to reduced comprehension time and decreased risk of misunderstanding. Simple structures may not be as beneficial to some readers, i.e., those with poorer background knowledge, as they are to others. Again, this appears to have significant implications for requirements guidelines, where, for example, the majority of use case description guidelines advocate extremely simple structures at all times. The problem being, that this may not always be beneficial to comprehension.

4.0 The Role of Background Knowledge

References to the importance of the reader's background knowledge occur throughout the discourse process literature [30]. Gordon and Scearce [20] argue that knowledge-based processes play a very early role in the processing of language while Kintsch and van Dijk [31] note that "comprehension always involves knowledge use and inference processes" and that the reader's knowledge largely determines the meaning derived from a text.

Graesser et al. [37] describe the knowledge-based inferences which readers generate during narrative comprehension as "critical building blocks", and suggest that the higher the number of levels of representation constructed and the higher the number of inferences generated by the reader at each level, the greater the improvement in comprehension. However, the generation of novel knowledge-based inferences is believed to place more demands on working memory. The greater the demands that a potential inference places on working memory, the lower the likelihood that it will be generated during comprehension. Thus, the more familiar the relevant background knowledge is to the reader, the more likely he or she is to generate inferences from it. Again, this would seem to suggest that background knowledge of the subject area is extremely important to the software engineer; in understanding the item under scrutiny, a view that is borne out by much software engineering literature (e.g., Curtis et al. [32, 33]).

4.1 Coherence and Background Knowledge

The interaction between the background knowledge of the reader and the coherence of the text is investigated by McNamara and Kintsch [34]. Their findings show that where recall of the text is concerned, i.e., just shallow understanding, coherence benefits readers' understanding, at least on a superficial level, but that domain knowledge can compensate for lack of coherence. They attribute this contradiction, between increased coherence both facilitating

comprehension and inhibiting understanding and learning, to the different levels which exist in the mental representation of the text generated by the reader. Readers are able to form a good textbase if the text is fully coherent. However, if the text lacks coherence at both local and global levels, readers are forced to fill the gaps themselves with a bridging inference. If they are unable to do so, they may be forced to build a situation model. This prevents them from basing their understanding purely on the superficial textbase. As a result, it enables them to be particularly successful in answering questions which require bridging inferences or problem solving and in a sorting task. Coherence and domain knowledge can, therefore, have different effects on understanding, depending on the different level of mental representation being constructed, and on the different nature of the tasks being performed by the reader.

Although increasing the coherence of a text will benefit readers with poor background knowledge, it may actually disadvantage readers with good background knowledge since it may reduce the amount of active processing that they perform during comprehension. Studies have shown that, generally, people are more likely to remember information which they have actively generated themselves than information which is simply presented to them without effort on their part and that they are more able to use information which they have generated themselves in new situations. Mannes and Kintsch [35] argue that the active generation of inferences during reading can be beneficial, if the inferences are successful, since it results in the creation of more links between the new information and the information which already resides in the reader's own knowledge base. They associate more active processing with the construction of a more detailed situation model and with superior performance on a problem solving task. Therefore, McNamara et al. [34] believe that it is sometimes better to allow readers to construct their own coherence rather than let the writer construct it for them in the text, although they add that it may be necessary to supply exactly the right amount of difficulty in a text to enable greater understanding by readers with high knowledge, making the text neither too easy nor too difficult. The degree of domain knowledge possessed by the reader is important. That is, the same text cannot be equally suitable for every reader.

5.0 The Reader in Comprehension

The lessons which may be learnt from discourse process theories are also affected by the sheer complexity of the process, or rather processes, of comprehension and by the differences between individual readers [36]. For example, research suggests that the selection of propositions which remain in short term memory (which approximately holds the most recent clause being comprehended) at the end of a sentence is influenced by the reader's goals [4]. Kintsch and van Dijk [28] also comment on different readers setting up different goals and proposing that the special purpose of the reader overrides the text structure. According to Van den Broek et al. [5], many researchers share the view that readers set for themselves a goal or standard for coherence. They then generate the amount of inferences necessary to attain this standard.

Thus, the discourse literature strongly indicates that the reader, as well as the text, is an important factor in understanding. The success of comprehension depends upon the extent to which there is harmony among the author's intended meaning of the text, the explicit text, and the reader's constructed meaning of the text [3].

6.0 Limitations of Discourse Process

Discourse process is a comparatively young field [37], and its researchers have proposed various theories which are often in opposition to one another [24]. Moreover, when theories and models are tested, there is a strong tendency for the results to "suggest" or "indicate" findings rather than unequivocally validate them. Graesser et al. [37], for example, discuss the need or the inevitability of integrating neuroscience into discourse process research and note that a few studies have already made use of functional brain imaging techniques to examine the activity of the brain during discourse comprehension. Furthermore, the traditional assumption which underlies many discourse process theories - that the reader constructs networks of propositions during comprehension - is now challenged: Stanfield and Zwaan [38] and Zwaan et al. [39], among others, proposing instead that during language comprehension people activate perceptual symbols of referents. In addition, the theories indicate that there is often a trade-off between the potential benefits and disadvantages of certain linguistic forms, such as complex structures which can result in both increased reading time and risk of misunderstanding and decreased reading time and risk of misunderstanding, depending on the circumstances.

Nevertheless, there are still significant lessons to be learned from discourse process which have important implications for the writing requirements. These include the importance of the reader's background knowledge, the existence of different levels of understanding, the need to make the reader do active work, the avoidance of pronouns and passive voice, and so on, all of which appear to have direct implications for the documents produced within the requirements phase.

7.0 Conclusions

The field of discourse process if often used as a theoretical underpinning for assertions about the nature of comprehension and it's implications for software engineering. However, few authors have attempted to unravel the complexities of comprehension before moving to its implications for software quality.

In contrast, this paper has sought to examine carefully the literature and lessons learned from discourse process in order to understand more fully the implications, particularly for the requirements phase.

The authors have found that there are many implications for the production of requirements, or indeed, any software engineering documentation. For example, perhaps the most intriguing finding is that despite the drive for ease of

comprehension it appears that it is not always beneficial for the reader if the author makes the text as explicit and coherent as possible. That is, a deeper understanding can be engendered where the reader needs to carry out work themselves, and construct their own explanations rather than for the text to present them with everything too easily. This is particularly so where the reader has a high level of background knowledge and is thus able to provide suitable explanations. This implies that a text should be challenging enough to encourage active processing by the reader but not so difficult as to become incomprehensible. Similar findings relate to the simplicity of text, the use of pronouns, and the use of the passive voice. For example, while there may be some gains in the speed in which very simple text, or text with pronouns, can be parsed such advantages are outweighed by the increased understanding, or lack of ambiguity gained by not using pronouns, or by having text that is neither too simple nor overly complex.

Clearly, this paper cannot, nor did it set out to, prescribe the ideal levels of complexity or structure for requirements documents. However, we believe that, particularly in highlighting the complex nature of text comprehension, and deriving clear implications, such lessons learned lay firmer foundations for the derivation of guidance, such as description guidelines for requirements authors. Hence, our hope is that such insights will prove beneficial not only to the quality of software documentation but also to the level of understanding gained by the breadth of stakeholders in the requirements phase.

References

1. Glass, R., 1998. *Software Runaways*. Harlow: Prentice Hall.
2. Bray, I.K., 2002. *An introduction to requirements engineering*. Harlow: Addison-Wesley.
3. Grasser, A.C., Singer, M. and Trabasso, T., 1994. Constructing inferences during narrative text comprehension. *Psychological Review*, 101 (3), 371-395.
4. Fletcher, C.R., Van Den Broek, P. and Arthur, E.J, 1996. A model of narrative comprehension and recall. in: B.K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, 141-163.
5. Van Den Broek, P., Risden, K., Fletcher, C.R. and Thurlow, R, 1996. A "landscape" view of reading: fluctuating patterns of activation and the construction of a stable memory representation. B.K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, 165-187.
6. Bamberg, M. and Moissinac, L., 2003. Discourse development. B.K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, 141-163.
7. Crystal D., 1997. *The Cambridge encyclopedia of language*. 2nd ed. Cambridge: Cambridge University Press.
8. Graesser, A.C., Millis, K.K. and Zwaan, R.A., 1997. Discourse comprehension. *Annual Review of Psychology*, 48, 163-189.
9. Cockburn, A., 2001. *Writing effective use cases*. London: Addison-Wesley.

10. Kulak, D. and Guiney, E., 2000. *Use cases: requirements in context*. London: ACM Press.
11. Foltz, P.W., 2003. Quantitative cognitive models of text and discourse processing. in: A.C. Grasser, M.A. Gernsbacher and S.R. Goldman, eds. *Handbook of discourse processes*. Mahwah, NJ: Lawrence Erlbaum Associates, 487-523.
12. Zwaan, R.A. and Singer, M., 2003. Text comprehension. in: A.C. Grasser, M.A. Gernsbacher and S.R. Goldman, eds. *Handbook of discourse processes*. Mahwah, NJ: Lawrence Erlbaum Associates, 83-121.
13. Gordon, P.C., Grosz, B.J. and Gilliom, L.A., 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17, 311-347.
14. Gibbs, R.W., 1996. Metaphor as a constraint on text understanding. B.K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, 215-240.
15. McNamara, D.S., Kintsch, E., Songer, N.B. and D Kintsch, W., 1996. Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14 (1), 1-43.
16. Britton, B.K. and Gulgoz, S., 1991. Using Kintsch's computational model to improve instructional text: effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83 (3), 329-345.
17. Gernsbacher, M.A., 1996. The structure-building framework: what it is, what it might also be, and why. .K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, 289-311.
18. Zwaan, R.A. and Singer, M, 2003. Text comprehension. A.C. Grasser, M.A. Gernsbacher and S.R. Goldman, eds. *Handbook of discourse processes*. Mahwah, NJ: Lawrence Erlbaum Associates, 83-121.
19. Ferreira, F., 2003. The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164-203.
20. Gordon, P.C. and Scearce, K.A., 1995. Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition*, 23 (3), 313-323.
21. Greene, S.B., McKoon, G. and Ratcliff, R., 1992. Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18 (2), 266-283.
22. McKoon, G., Gerrig, R.J. and Greene, S.B., 1996. Pronoun resolution without pronouns: some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22 (4), 919-932.
23. Gordon, P.C. and Chan, D., 1995. Pronouns, passives, and discourse coherence. *Journal of Memory and Language*, 34, 216-231.
24. Gordon, P.C., Grosz, B.J. and Gilliom, L.A., 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17, 311-347.
25. Harley, T.A., 2001. *The psychology of language: From data to theory*. 2nd ed. Hove: Psychology Press.
26. Johnson-Laird, P.N., 1968. The choice of the passive voice in a communicative task. *British Journal of Psychology*, 59, 7-15.

27. Ferreira, F., 1994. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33 (6), 715-736.
28. Kintsch, W. and Van Dijk, T.A., 1978. Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
29. Moore, J.D. and Wiemer-Hastings, P., 2003. Discourse in computational linguistics and artificial intelligence. A.C. Grasser, M.A. Gernsbacher and S.R. Goldman, eds. *Handbook of discourse processes*, Mahwah, NJ: Lawrence Erlbaum Associates, 439-485.
30. Mannes, S. and St. George, M., 1996. Effects of prior knowledge on text comprehension: a simple modeling approach. in: B.K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates, 115-139.
31. Van Dijk, T.A. and Kintsch, W: Strategies of Discourse Comprehension: Academic Press New York, 1983
32. Curtis, B., H. Krasner, et al. (1988). A field study of the software design process for large systems. Communications of the ACM, 31(11): 1268-1287.
33. Curtis, B., D. Walz, et al. (1989). Studying the Process of Software Design Teams. Proceedings of the 5th International Software Process Workshop, Kennebunkport, Maine, USA, IEEE Computer Society Press.
34. McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. Discourse Processes, 22, 247-287
35. Mannes, S.M. and Kintsch, W., 1987. Knowledge organization and text organization. *Cognition and Instruction*, 4 (2), 91-115.
36. Goldman, S.R., Varma, S., and Cote, N., 1996. Extending capacity-constrained construction integration: toward "smarter" and flexible models of text comprehension. in: B.K. Britton and A.C. Graesser, eds. *Models of understanding text*. Mahwah, NJ: Lawrence Erlbaum Associates
37. Graesser, A.C., Gernsbacher, M.A., and Goldman, S.R., 2003. Introduction to the handbook of discourse processes. A.C. Grasser, M.A. Gernsbacher and S.R. Goldman, eds. *Handbook of discourse processes*, Mahwah, NJ: Lawrence Erlbaum Associates.
38. Stanfield R.A. and Zwaan, R.A., 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12 (2), 153-156.
39. Zwaan, R.A., Stanfield, R.A. and Yaxley, R.H., 2002. Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13 (2), 168-171.